```
OM of: US-09-126-945B-2_COPY_236_243 to: EST:*   out_format : pfs

Date: Sep 7, 2001  6:00 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlh
 -Q=/cgn2_1/USPTO_spool/US09126945/runat_07092001_152108_3642/app_query.fasta_1.487
 -DB=EST -QFMT=fastap -SUFFIX=olig.rst -GAPOP=4.500 -GAPEXT=0.050
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=60.000 -XGAPEXT=60.000 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=60.000 -YGAPEXT=60.000 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15
 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=2000000000
 -USER=US09126945_@CGN1_1_4183 -NCPU=6 -ICPU=3 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-126-945B-2_COPY_236_243
Query length: 8
Database: EST:*
Database sequences: 10228115
Database length: 431459454
Search time (sec): 1300.560000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000
WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000


score_list:
Sequence              Strd Orig    ZScore    EScore Len  ! Documentation ..
gb_est78:BE757615      +     8.00  168.44     2.90    539  ! BE757615 212104 MARC 2BOV Bos t
gb_est42:AW101885      -     7.00  149.36    33.55    391  ! AW101885 sd72a05.y1 Gm-c1008 Gl
gb_est80:BE841998      -     7.00  149.10    34.69    408  ! BE841998 MR4-ST0098-280600-010-
gb_est97:BG159088      +     7.00  147.29    43.76    549  ! BG159088 RHIZ2_17_F11.b1_A003 R
gb_est16:AI119983      -     7.00  147.22    44.13    555  ! AI119983 uc25a01.r1 Soares_mamm
gb_gss12:AQ510431      +     7.00  147.13    44.63    563  ! AQ510431 nbxb0095M16r CUGI Rice
gb_est27:AJ393331      -     7.00  144.97    58.92    803  ! AJ393331 AJ393331 dkfz426 Gallu
gb_est92:BF795883      +     7.00  144.16    65.36    917  ! BF795883 602259125F1 NIH_MGC_85
gb_est81:BE978811      -     6.00  138.23   139.82     77  ! BE978811 bs83b08.y1 Drosophila m
gb_est111:Z31199       +     6.00  134.45   226.91    143  ! Z31199 MMTEST592 Mouse testis T
gb_est81:BE938486      +     6.00  133.66   251.38    163  ! BE938486 RC0-TN0079-290800-023-
gb_gss35:B42379        -     6.00  133.51   256.19    167  ! B42379 HS-1055-B2-C09-MF.abi C1
gb_est54:BB019091      +     6.00  133.12   269.30    178  ! BB019091 BB019091 RIKEN full-le
gb_gss18:AQ925777      +     6.00  132.95   275.20    183  ! AQ925777 RPCI-23-297P13.TV RPCI
gb_est32:AV057324      +     6.00  132.82   279.89    187  ! AV057324 AV057324 Mus musculus
gb_est58:BB153887      +     6.00  132.59   288.05    194  ! BB153887 BB153887 RIKEN full-le
gb_est2:AA091451       -     6.00  131.56   329.08    230  ! AA091451 il2866.seq.F Human fet
gb_est52:AW838681      -     6.00  131.48   332.43    233  ! AW838681 RC4-LT0055-130200-011-
gb_est34:AV174026      -     6.00  131.40   335.77    236  ! AV174026 AV174026 Mus musculus
gb_est64:BB431289      +     6.00  131.40   335.77    236  ! BB431289 BB431289 RIKEN full-le
gb_est103:C12830       +     6.00  131.30   340.22    240  ! C12830 C12830 Yuji Kohara unpub
gb_est36:AV282217      +     6.00  131.22   343.54    243  ! AV282217 AV282217 RIKEN full-le
gb_est54:BB016278      +     6.00  131.20   344.64    244  ! BB016278 BB016278 RIKEN full-le
gb_est69:BE059133      -     6.00  131.10   349.06    248  ! BE059133 sn25g02.y1 Gm-c1016 Gl
gb_est35:AV263082      +     6.00  131.07   350.16    249  ! AV263082 AV263082 RIKEN full-le
gb_est35:AV262061      +     6.00  131.05   351.26    250  ! AV262061 AV262061 RIKEN full-le
gb_est53:AW885773      -     6.00  131.02   352.36    251  ! AW885773 RC4-OT0071-050500-015-
gb_est68:BB598201      -     6.00  131.02   352.36    251  ! BB598201 BB598201 RIKEN full-le
gb_est68:BB576695      -     6.00  130.88   358.93    257  ! BB576695 BB576695 RIKEN full-le
gb_est4:AA245510       -     6.00  130.81   362.20    260  ! AA245510 my51h12.r1 Barstead mo
gb_est96:BG077091      -     6.00  130.81   362.20    260  ! BG077091 H3010F03-5 NIA Mouse 1
gb_est75:BE530105      -     6.00  130.79   363.29    261  ! BE530105 M76D14STM Arabidopsis
gb_est50:AW714267      +     6.00  130.53   375.21    272  ! AW714267 h7b01ne.f1 Neurospo
gb_gss33:AZ860786      +     6.00  130.49   377.37    274  ! AZ860786 2M0166N21R Mouse 10
gb_est9:AA563231       -     6.00  130.44   379.52    276  ! AA563231 v165g05.r1 Knowles
gb_est92:BF799425      -     6.00  130.40   381.67    278  ! BF799425 RC3-C10042-091000-0
gb_est34:AV209286      +     6.00  130.38   382.74    279  ! AV209286 AV209286 RIKEN full
gb_est62:BB342044      +     6.00  130.38   382.74    279  ! BB342044 BB342044 RIKEN full
gb_est50:AW714267      +     6.00  130.53   375.21    272  ! AW714267 h7b01ne.f1 Neurospo
gb_gss33:AZ860786      +     6.00  130.49   377.37    274  ! AZ860786 2M0166N21R Mouse 10
gb_est9:AA563231       -     6.00  130.44   379.52    276  ! AA563231 v165g05.r1 Knowles
gb_est92:BF799425      -     6.00  130.40   381.67    278  ! BF799425 RC3-C10042-091000-0
gb_est34:AV209286      +     6.00  130.38   382.74    279  ! AV209286 AV209286 RIKEN full
gb_est62:BB342044      +     6.00  130.38   382.74    279  ! BB342044 BB342044 RIKEN full

seq_name: gb_est78:BE757615

seq_documentation_block:
LOCUS       BE757615      539 bp    mRNA            EST       15-SEP-2000
DEFINITION  212104 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BE757615
VERSION     BE757615.1  GI:10171607
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 539)
  AUTHORS   Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,
            Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Laegreid
            ,W.W. and Keele,J.W.
  TITLE     Design and use of four pooled tissue normalized cDNA libraries for
            EST discovery in cattle
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 64  row: O  column: 13
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="MARC 2BOV"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled tissue from testis, thymus,
                     semitendonosus muscle, longissimus muscle, pancreas,
                     adrenal, and endometrium."
BASE COUNT        98 a    168 c    193 g     80 t
ORIGIN


alignment_scores:
            Quality:    8.00              Length:      8
              Ratio:   1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x BE757615   ..

 Align seg 1/1  to: BE757615  from: 1  to: 539

       1 ThrAspSerGluValAspSerSer 8
         ||||||||||||||||||||||||
     458 ACGGACAGCGAGGTGGACTCGTCC 481

seq_name: gb_est42:AW101885

seq_documentation_block:
```

```
LOCUS       AW101885      391 bp    mRNA            EST       18-APR-2000
DEFINITION  sd72a05.y1 Gm-c1008 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1008-1833 5' similar to SW:TBA1_ANEPH P33623 TUBULIN ALPHA-1
            CHAIN. [1] ;, mRNA sequence.
ACCESSION   AW101885
VERSION     AW101885.1  GI:6072498
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
REFERENCE   1  (bases 1 to 391)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: Genome Systems, Inc. 4633 World
            Parkway Circle St. Louis, Missouri 63134 For further information
            call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
            427-3324 or contact: clones@genomesystems.com or
            info@genomesystems.com web site: www.genomesystems.com
            Putative full length read
            vector to vector length is 392
            Insert Length: 542   Std Error: 0.00
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .391
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1008-1833"
                     /clone_lib="Gm-c1008"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
                     cDNA library was constructed from mRNA isolated from whole
                     young pods, approximately 2cm long, of 12-week-old
                     greenhouse grown plants. The library was prepared using
                     the Life Technologies pSuperScript cDNA library
                     construction kit. Complementary DNA was synthesized from
                     mRNA using a poly (dT) sequence with a NotI restriction
                     site. SalI linkers adapters were ligated to the
                     blunt-ended cDNA fragments followed by NotI digestion. The
                     cDNA fragments were directionally cloned into the
                     NotI-SalI restriction site of the pSPORT1 vector. The
                     ligated cDNA fragments were transformed into E. coli
                     ElectroMax DH10B host cells. This library was constructed
                     by Dr. Lila Vodkin and Dr. Anu Khanna."
BASE COUNT       87 a     69 c    108 g    127 t
ORIGIN


alignment_scores:
           Quality:    7.00            Length:      7
             Ratio:  1.000              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x AW101885/rev   ..

 Align seg 1/1  to reverse of: AW101885  from: 1  to: 391

      2 AspSerGluValAspSerSer 8
```

```
                  ||||||||||||||||||||
      234 GATTCTGAAGTAGATTCTAGT 214

 seq_name: gb_est80:BE841998

 seq_documentation_block:
 LOCUS       BE841998      408 bp    mRNA            EST       22-SEP-2000
 DEFINITION  MR4-ST0098-280600-010-c07 ST0098 Homo sapiens cDNA, mRNA sequence.
 ACCESSION   BE841998
 VERSION     BE841998.1  GI:10274376
 KEYWORDS    EST.
 SOURCE      human.
   ORGANISM  Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE   1  (bases 1 to 408)
   AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
             Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
             Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
             Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
             ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
             Simpson,A.J.
   TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
             sequence tags
   JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   MEDLINE   20202663
 COMMENT     Contact: Simpson A.J.G.
             Laboratory of Cancer Genetics
             Ludwig Institute for Cancer Research
             Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
             Brazil
             Tel: +55-11-2704922
             Fax: +55-11-2707001
             Email: asimpson@ludwig.org.br
             This sequence was derived from the FAPESP/LICR Human Cancer Genome
             Project. This entry can be seen in the following URL
             (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=&t2=MR4-ST0098-280
             600-010-c07&t3=2000-06-28&t4=1)
             Seq primer: puc 18 forward
             High quality sequence start: 37
             High quality sequence stop: 408.
 FEATURES             Location/Qualifiers
      source          1. .408
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone_lib="ST0098"
                      /dev_stage="Adult"
                      /note="Organ: stomach; Vector: puc18; Site_1: SmaI;
                      Site_2: SmaI; A mini-library was made by cloning products
                      derived from  ORESTES PCR (U.S. Letters Patent application
                      No. 196,716 -  Ludwig Institute for Cancer Research)
                      profiles into  the pUC 18 vector. Reverse transcription of
                      tissue mRNA and cDNA amplification were performed under
                      low stringency conditions."
 BASE COUNT      159 a     66 c     76 g    107 t
 ORIGIN


 alignment_scores:
            Quality:    7.00            Length:      7
              Ratio:  1.000              Gaps:      0
  Percent Similarity: 100.000   Percent Identity: 100.000

 alignment_block:
  US-09-126-945B-2_COPY_236_243 x BE841998/rev   ..

  Align seg 1/1  to reverse of: BE841998  from: 1  to: 408

       2 AspSerGluValAspSerSer 8
         ||||||||||||||||||||
     308 GATTCTGAAGTGGACTCATCC 288

 seq_name: gb_est97:BG159088
```

```
seq_documentation_block:
LOCUS       BG159088      549 bp    mRNA            EST       06-FEB-2001
DEFINITION  RHIZ2_17_F11.b1_A003 Rhizome2 (RHIZ2) Sorghum propinquum cDNA, mRNA
            sequence.
ACCESSION   BG159088
VERSION     BG159088.1  GI:12692752
KEYWORDS    EST.
SOURCE      Sorghum propinquum.
  ORGANISM  Sorghum propinquum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 549)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Paterson,A., Sudman,M. and Pratt
            ,L.H.
  TITLE     An EST database from Sorghum: Sorghum propinquum rhizomes
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cordonnier-Pratt MM
            Department of Botany
            The University of Georgia
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 542 1805
            Email: mmpratt@uga.edu
            Seq primer: JEN REV
            High quality sequence stop: 486
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1..549
                     /organism="Sorghum propinquum"
                     /db_xref="taxon:132711"
                     /clone_lib="Rhizome2 (RHIZ2)"
                     /note="Organ: Rhizomes; Vector: pBluescript II from Lambda
                     Zap II; Site_1: XhoI; Site_2: EcoRI; The library was made
                     from poly-A RNA in the cloning vector lambda ZAP II.
                     Clones to be sequenced were prepared by mass excision."
BASE COUNT      130 a    133 c    156 g    130 t
ORIGIN


alignment_scores:
           Quality:    7.00            Length:      7
             Ratio:  1.000              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x BG159088   ..

 Align seg 1/1  to: BG159088  from: 1  to: 549

       1 ThrAspSerGluValAspSer 7
         |||||||||||||||||||||
     523 ACTGACAGTGAAGTTGACTCC 543

seq_name: gb_est16:AI119983

seq_documentation_block:
LOCUS       AI119983      555 bp    mRNA            EST       02-SEP-1998
DEFINITION  uc25a01.r1 Soares_mammary_gland_NbMMG Mus musculus cDNA clone
            IMAGE:1398984 5' similar to gb:J04982_cds1 ADP,ATP CARRIER PROTEIN,
            HEART/SKELETAL MUSCLE ISOFORM (HUMAN); gb:U27315 Mus musculus
            adenine nucleotide translocase-1 (MOUSE);, mRNA sequence.
ACCESSION   AI119983
VERSION     AI119983.1  GI:3520307
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 555)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:910700
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 258.
FEATURES             Location/Qualifiers
     source          1..555
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1398984"
                     /clone_lib="Soares_mammary_gland_NbMMG"
                     /sex="male"
                     /tissue_type="mammary gland"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia
                     ) with a modified polylinker; Site_1: Not I; Site_2: Eco
                     RI; 1st strand cDNA was primed with a Not I - oligo(dT)
                     primer (5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGAATGGTTTTTTTTTTTTTTTTTTTT
                     T 3'); double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
                     constructed and normalized by Bento Soares and M.Fatima
                     Bonaldo."
BASE COUNT      155 a    101 c    151 g    147 t      1 others
ORIGIN


alignment_scores:
           Quality:    7.00            Length:      7
             Ratio:  1.000              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x AI119983/rev   ..

 Align seg 1/1  to reverse of: AI119983  from: 1  to: 555

       2 AspSerGluValAspSerSer 8
         ||||||||||||||||||||
      50 GACTCCGAAGTAGACAGCTCT 30

seq_name: gb_gss12:AQ510431

seq_documentation_block:
LOCUS       AQ510431      563 bp    DNA             GSS       04-MAY-1999
DEFINITION  nbxb0095M16r CUGI Rice BAC Library Oryza sativa genomic clone
            nbxb0095M16r, DNA sequence.
ACCESSION   AQ510431
VERSION     AQ510431.1  GI:4733035
KEYWORDS    GSS.
SOURCE      Oryza sativa.
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 563)
  AUTHORS   Wing,R.A. and Dean,R.A.
```

```
  TITLE     A BAC End Sequencing Framework to Sequence the Rice Genome
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: GGAAACAGCTATGACCATG
            Class: BAC ends
            High quality sequence stop: 411.
FEATURES             Location/Qualifiers
     source          1. .563
                     /organism="Oryza sativa"
                     /strain="Japonica"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:4530"
                     /clone="nbxb0095M16r"
                     /clone_lib="CUGI Rice BAC Library"
                     /tissue_type="Leaf"
                     /lab_host="E. coli DH10B"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII; Rice is one of two most popular grains in the
                     world. Half of the world  population especially those
                     inhabiting highly populated areas of the  humid tropics
                     and subtropics, rely on rice as their primary source of
                     carbohydrate. Monocotyledonous rice is a diploid plant
                     (2n=24) with a  haploid genome equivalent of 431 Mbp
                     (Arumuganathan and Earle, 1991).  The relatively small
                     genome of rice, three times larger than that of
                     Arabidopsis, makes it suitable for genomic studies. In
                     order to  facilitate positional cloning, physical mapping
                     and genome sequencing  of rice, we have constructed a BAC
                     library from Oryza sativa, Nipponbare  variety. The
                     library contains 36,864 clones with an average insert size
                     of 128.5 Kb providing 10.9 haploid genome equivalents. The
                     deep  coverage allows the isolation a particular sequence
                     with a probability  of 99.9 %. Two high density filters,
                     each containing 18,432 clones  (doubly spotted), represent
                     the whole library for colony screening."
BASE COUNT      179 a     97 c     90 g    197 t
ORIGIN


alignment_scores:
            Quality:    7.00             Length:      7
              Ratio:   1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x AQ510431   ..

 Align seg 1/1  to: AQ510431  from: 1  to: 563

       2 AspSerGluValAspSerSer 8
         ||||||||||||||||||||
     137 GATTCGGAAGTTGACTCATCC 157

seq_name: gb_est27:AJ393331

seq_documentation_block:
LOCUS       AJ393331      803 bp    mRNA            EST       25-JAN-2001
DEFINITION  AJ393331 dkfz426 Gallus gallus cDNA clone 15c3r1, mRNA sequence.
ACCESSION   AJ393331
VERSION     AJ393331.1  GI:7122087
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 803)
  AUTHORS   Abdrakhmanov,I., Lodygin,D., Geroth,P., Arakawa,H., Law,A., Plachy
            ,J., Korn,B. and Buerstedde,J.M.
  TITLE     A large database of chicken bursal ESTs as a resource for the
            analysis of vertebrate gene function
  JOURNAL   Genome Res. 10 (12), 2062-2069 (2000)
  MEDLINE   20568495
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinistr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES             Location/Qualifiers
     source          1. .803
                     /organism="Gallus gallus"
                     /strain="CB"
                     /db_xref="taxon:9031"
                     /clone="15c3r1"
                     /clone_lib="dkfz426"
                     /tissue_type="Bursa of Fabricius"
BASE COUNT      265 a    155 c    186 g    194 t      3 others
ORIGIN


alignment_scores:
            Quality:    7.00             Length:      7
              Ratio:   1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x AJ393331/rev   ..

 Align seg 1/1  to reverse of: AJ393331  from: 1  to: 803

       1 ThrAspSerGluValAspSer 7
         |||||||||||||||||||||
     587 ACAGATTCTGAAGTTGACTCC 567

seq_name: gb_est92:BF795883

seq_documentation_block:
LOCUS       BF795883      917 bp    mRNA            EST       12-JAN-2001
DEFINITION  602259125F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4342183 5',
            mRNA sequence.
ACCESSION   BF795883
VERSION     BF795883.1  GI:12100937
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 917)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis Staudt, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9957  row: b  column: 08
            High quality sequence stop: 633.
FEATURES             Location/Qualifiers
     source          1. .917
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4342183"
                     /clone_lib="NIH_MGC_85"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
```

```
                /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                Average insert size 1.867 kb. Library enriched for
                full-length clones and constructed by Life Technologies.
                Note: this is a NIH_MGC Library."
BASE COUNT      215 a    229 c    281 g    191 t      1 others
ORIGIN


alignment_scores:
            Quality:    7.00             Length:      7
              Ratio:  1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x BF795883   ..

 Align seg 1/1  to: BF795883  from: 1  to: 917

       2 AspSerGluValAspSerSer 8
         ||||||||||||||||||||
     711 GACTCCGAAGTTGATTCCTCT 731

seq_name: gb_est81:BE978811

seq_documentation_block:
LOCUS       BE978811       77 bp    mRNA            EST       04-OCT-2000
DEFINITION  bs83b08.y1 Drosophila melanogaster adult testis library Drosophila
            melanogaster cDNA clone bs83b08 5', mRNA sequence.
ACCESSION   BE978811
VERSION     BE978811.1  GI:10610669
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 77)
  AUTHORS   Andrews, J., Bouffard, G. and Oliver, B.
  TITLE     Drosophila melanogaster testis expressed sequence tags
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Brian Oliver
            Laboratory of Cellular and Developmental Biology
            NIDDK, National Institutes of Health
            6 Center Drive MSC 2715, Bldg 6, Rm B1-13, Bethesda, MD 20892 USA
            Fax: (301) 496 5239
            Email: oliver@helix.nih.gov,
            http://www.niddk.nih.gov/intram/people/boliver.htm
            Tissue isolation and library construction performed at the National
            Institute of Diabetes and Digestive and Kidney Diseases, NIH (see
            http://www.niddk.nih.gov/intram/people/boliver.htm). DNA sequencing
            and analyses performed by National Institutes of Health Intramural
            Sequencing Center (NISC; see http://www.nisc.nih.gov).
            Plate: 83  row: b  column: 08
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES                Location/Qualifiers
     source             1. .77
                        /organism="Drosophila melanogaster"
                        /strain="y[*] w[67c1]/Y"
                        /db_xref="taxon:7227"
                        /clone="bs83b08"
                        /clone_lib="Drosophila melanogaster adult testis library"
                        /sex="male"
                        /dev_stage="1-5 day adult"
                        /lab_host="SOLR (Stratagene)"
                        /note="Organ: testis; Vector: pBlueScript SK (Stratagene);
                        Site_1: EcoR I; Site_2: Xho I; Testes dissected from 1-5
                        day adult y[*] w[67c1]/Y males raised at 25oC.  RNA
                        isolated using Trizol (Life Technologies) and a single
                        round of Poly(A)+ selection using Oligotex (Qiagen).  cDNA
                        library constructed using Stratagene ZAP-cDNA syntesis
                        kit.  Oligo dT-primed, size fractionated ~1-6 kb, and
                        directionally cloned at EcoRI and XhoI in Uni-ZAP XR.
                        Following a single round of amplification pBlueScript SK
                        phagemids were mass excised.  A distribution channel for
                        clones is being sought, but not currently available.
                        Requests for clones cannot be honored."
BASE COUNT      19 a     22 c     15 g     21 t
ORIGIN


alignment_scores:
            Quality:    6.00             Length:      6
              Ratio:  1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x BE978811/rev   ..

 Align seg 1/1  to reverse of: BE978811  from: 1  to: 77

       3 SerGluValAspSerSer 8
         ||||||||||||||||||
      45 TCGGAAGTAGATTCAAGT 28

seq_name: gb_est111:Z31199

seq_documentation_block:
LOCUS       Z31199         143 bp    mRNA            EST       24-JAN-1995
DEFINITION  MMTEST592 Mouse testis T-ZAP Mus musculus cDNA, mRNA sequence.
ACCESSION   Z31199
VERSION     Z31199.1  GI:633431
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 143)
  AUTHORS   Yuan,L., Liu,J.G. and Hoog,C.
  TITLE     Rapid cDNA sequencing in combination with RNA expression studies
            identifies a large number of male germ cell-specific sequence tags
  JOURNAL   Biol. Reprod. 52, 131-138 (1995)
  MEDLINE   95226560
COMMENT     Contact: C. Hoog
            Dept. of Molecular Genetics
            Karolinska Institutet
            Box 60400,10401 Stockholm,SWEDEN.
FEATURES                Location/Qualifiers
     source             1. .143
                        /organism="Mus musculus"
                        /strain="CBA"
                        /db_xref="taxon:10090"
                        /clone_lib="Mouse testis T-ZAP"
                        /note="Testis from adult mouse"
BASE COUNT      33 a     39 c     29 g     41 t      1 others
ORIGIN


alignment_scores:
            Quality:    6.00             Length:      6
              Ratio:  1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x Z31199   ..

 Align seg 1/1  to: Z31199  from: 1  to: 143

       3 SerGluValAspSerSer 8
         ||||||||||||||||||
      58 AGCGAAGTGGACTCCTCC 75

seq_name: gb_est81:BE938486

seq_documentation_block:
LOCUS       BE938486       163 bp    mRNA            EST       02-OCT-2000
```

```
DEFINITION  RC0-TN0079-290800-023-b06 TN0079 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE938486
VERSION     BE938486.1  GI:10465862
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 163)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=RC0-TN0079-290
            800-023-b06&t3=2000-08-29&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 43
            High quality sequence stop: 163.
FEATURES             Location/Qualifiers
     source          1. .163
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="TN0079"
                     /dev_stage="Adult"
                     /note="Organ: testis_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       38 a      43 c      43 g      39 t
ORIGIN


alignment_scores:
            Quality:    6.00             Length:      6
              Ratio:   1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x BE938486   ..

 Align seg 1/1  to: BE938486  from: 1  to: 163

       3 SerGluValAspSerSer 8
         ||||||||||||||||||
     19 TCAGAGGTAGACAGCAGC 36

seq_name: gb_gss35:B42379

seq_documentation_block:
LOCUS       B42379         167 bp    DNA             GSS       18-OCT-1997
DEFINITION  HS-1055-B2-C09-MF.abi CIT Human Genomic Sperm Library C Homo
            sapiens genomic clone Plate=CT 777 Col=18 Row=F, DNA sequence.
ACCESSION   B42379
VERSION     B42379.1  GI:2546631
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 167)
  AUTHORS   Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,
            Tracicoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E.
  TITLE     Construction of a Characterized Clone Resource for Genomic
            Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
            Tagged Connectors
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mahairas GG, Zackrone KD, Hood L
            University of Washington
            Seattle, WA 98195, USA
            Tel: (206) 616-8744
            Fax: (206) 685-7301
            Email: kzackron@u.washington.edu
            Sequence Tagged Connector
            Plate: CT 777  row: F  column: 18
            Class: BAC ends
            High quality sequence stop: 167.
FEATURES             Location/Qualifiers
     source          1. .167
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=CT 777 Col=18 Row=F"
                     /clone_lib="CIT Human Genomic Sperm Library C"
                     /sex="M"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
BASE COUNT       43 a      40 c      40 g      43 t       1 others
ORIGIN


alignment_scores:
            Quality:    6.00             Length:      6
              Ratio:   1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x B42379/rev   ..

 Align seg 1/1  to reverse of: B42379  from: 1  to: 167

       3 SerGluValAspSerSer 8
         ||||||||||||||||||
     96 TCAGAAGTTGATTCCAGT 79

seq_name: gb_est54:BB019091

seq_documentation_block:
LOCUS       BB019091       178 bp    mRNA            EST       22-JUN-2000
DEFINITION  BB019091 RIKEN full-length enriched, adult male testis (DH10B) Mus
            musculus cDNA clone 4930583H10 3', mRNA sequence.
ACCESSION   BB019091
VERSION     BB019091.1  GI:8190850
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 178)
  AUTHORS   Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci
            ,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
            Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
            Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
            Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.
            , Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.
            , Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata
            ,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y.
            , Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya
            ,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I.
```

```
          , Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino
          ,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Konno,H., et al.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Yoshihide Hayashizaki
          Genome Exploration Research Group, Life Science Tsukuba Center,
          Genome Science Laboratory
          The Institute of Physical and Chemical Research (RIKEN), Genomic
          Sciences Center
          3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
          Tel: +81-298-36-9013
          Fax: +81-298-36-9098
          Email: genome-res@rtc.riken.go.jp,
          URL:http://genome.rtc.riken.go.jp/
          Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki
          ,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
          Thermostabilization and thermoactivation of thermolabile enzymes by
          trehalose and its application for the synthesis of full length
          cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
           Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
          Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
          ,Y. and Hayashizaki,Y.
           Automated filtration-based high-throughput plasmid preparation
          system. Genome Res. 9 (5), 463-470 (1999)
           Carninci,P. and Hayashizaki,Y.
           High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
          19-44 (1999)
           Please visit our web site (http://genome.rtc.riken.go.jp) for
          further details.
FEATURES             Location/Qualifiers
     source          1..178
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="4930583H10"
                     /clone_lib="RIKEN full-length enriched, adult male testis
                     (DH10B)"
                     /sex="male"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3']. cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. Second strand cDNA was prepared with the
                     primer adapter of sequence [5'
                     GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA
                     was cloned into the XhoI and BamHI sites. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI."
BASE COUNT       54 a     32 c     59 g     33 t
ORIGIN

.

alignment_scores:
           Quality:    6.00             Length:      6
             Ratio:  1.000               Gaps:       0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x BB019091   ..

 Align seg 1/1  to: BB019091  from: 1  to: 178

       3 SerGluValAspSerSer 8
         ||||||||||||||||||
```

```
      40 AGCGAGGTGGACAGCAGT 57

seq_name: gb_gss18:AQ925777

seq_documentation_block:
LOCUS       AQ925777       183 bp    DNA             GSS       21-DEC-1999
DEFINITION  RPCI-23-297P13.TV RPCI-23 Mus musculus genomic clone RPCI-23-297P13
            , DNA sequence.
ACCESSION   AQ925777
VERSION     AQ925777.1  GI:6614697
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 183)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 297  row: P  column: 13
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..183
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-297P13"
                     /clone_lib="RPCI-23"
                     /sex="Female"
                     /lab_host="DH10B"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies)."
BASE COUNT       66 a     39 c     32 g     46 t
ORIGIN


alignment_scores:
           Quality:    6.00             Length:      6
             Ratio:  1.000               Gaps:       0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x AQ925777   ..

 Align seg 1/1  to: AQ925777  from: 1  to: 183

       3 SerGluValAspSerSer 8
         ||||||||||||||||||
      42 TCTGAAGTCGACTCCAGC 59

seq_name: gb_est32:AV057324
```

```
seq_documentation_block:
LOCUS       AV057324       187 bp    mRNA            EST       23-JUN-1999
DEFINITION  AV057324 Mus musculus pancreas C57BL/6J adult Mus musculus cDNA
            clone 1810045G16, mRNA sequence.
ACCESSION   AV057324
VERSION     AV057324.1  GI:5157071
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 187)
  AUTHORS   Carninci,P., Shibata,K., Ozawa,Y., Konno,H., Itoh,M., Aizawa,K.,
            Akahira,S., Akiyama,J., Fukuda,S., Fukunishi,Y., Funayama,T., Hara
            ,A., Hayatsu,N., Hori,F., Ishikawa,T., Itoh,M., Izawa,M., Kawai,J.,
            Kikuchi,N., Kojima,Y., Matsuyama,T., Niitsuma,H., Oda,H., Owa,C.,
            Sato,K., Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara
            ,Y., Suzuki,H., Suzuki,H., Tateno,M., Tomaru,Y., Tominaga,N.,
            Watanabe,S., Yagame,M., Yamamura,T., Yokota,T., Yoshino,M.,
            Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Chie Owa
            Genome Science Laboratory
            RIKEN
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: 81-298-36-9145
            Fax: 81-298-36-9098
            Email: genome-res@rtc.riken.go.jp
            Thermostabilization and thermoactivation of thermolabile enzymes by
            trehalose and its application for the synthesis of full length cDNA
            (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
            Transcriptional sequencing: A method for DNA sequencing using RNA
            polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
            Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .187
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="1810045G16"
                     /clone_lib="Mus musculus pancreas C57BL/6J adult"
                     /sex="male"
                     /tissue_type="pancreas"
                     /dev_stage="adult"
BASE COUNT       31 a     49 c     39 g     68 t
ORIGIN


alignment_scores:
            Quality:    6.00             Length:      6
              Ratio:  1.000               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x AV057324   ..

 Align seg 1/1  to: AV057324  from: 1  to: 187

       3 SerGluValAspSerSer 8
         !!!!!!!!!!!!!!!!!!
       1 TCAGAGGTCGATTCTAGT 18
```